C-1027 chromophore ⟹ **Benzenoid diradical** ⟹ DNA cleavage
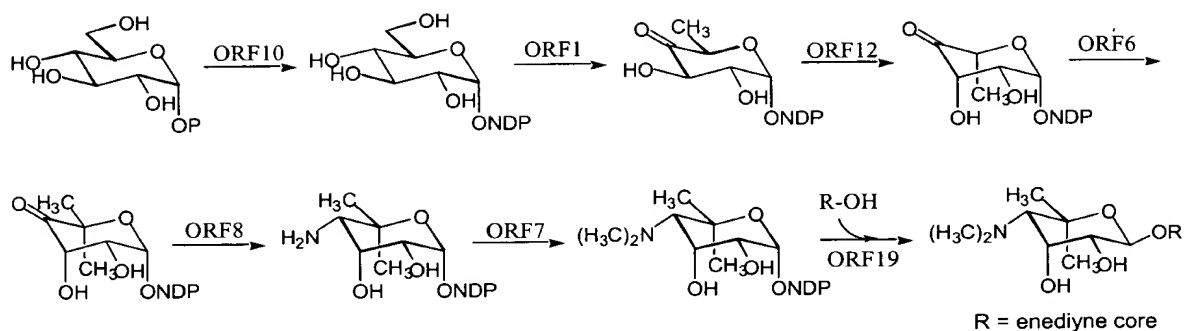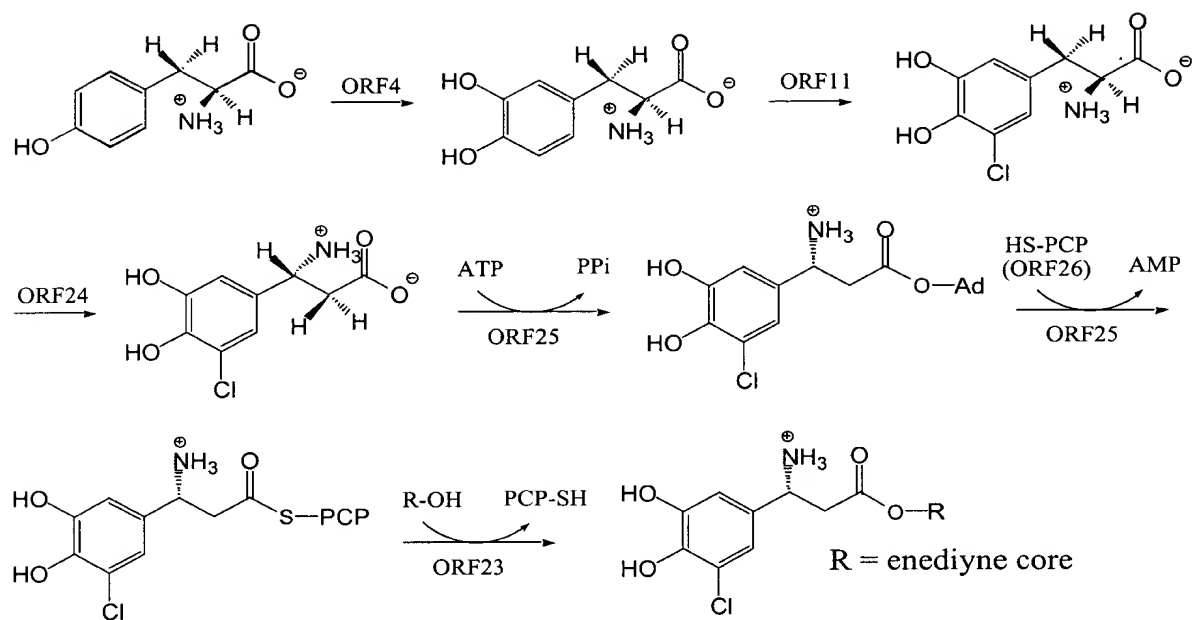
Fig. 1

Fig. 2

ORF10: dNDP-glucose synthase, 355 aa
ORF1: dNDP-glucose dehydratase, 332 aa
ORF12: epimerase, 192 aa
ORF8: aminotransferase, 410 aa

ORF6: C-methyltransferase, 423 aa
ORF7: N-methyltransferase, 244 aa
ORF19: glycosyl transferase, 459 aa

R = enediyne core

ORF4: Hydroxylase, 527 aa
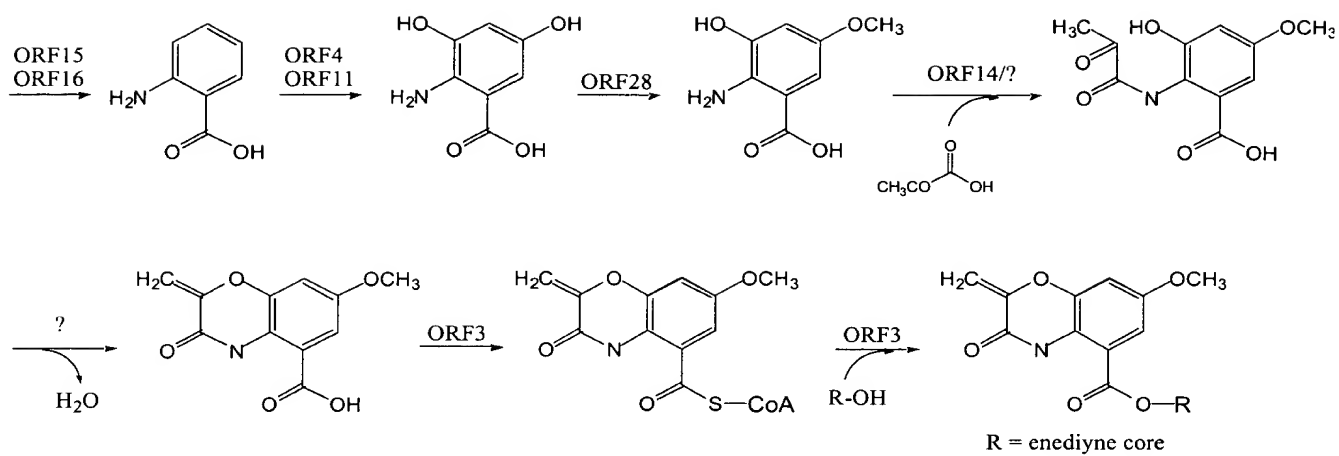ORF11: Hydroxylase/halogenase, 492/494 aa
ORF24: Aminomutase, 539 aa
ORF23: Type II NRPS condensation enzyme, 459 aa
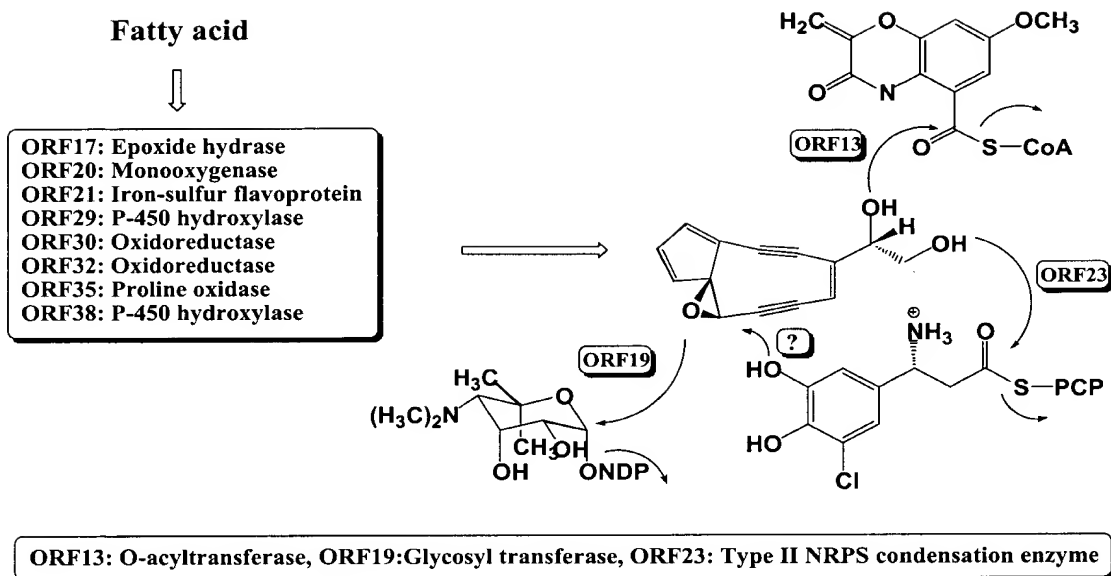ORF25: Type II NRPS adenylation enzyme, 716 aa
ORF26: Type II peptidyl carrier protein, 93 aa

Fig. 3A

ORF15: Anthranilate synthase I, 493 aa   ORF3: Coenzyme F390 synthetase, 463 aa
ORF16: Anthranilate synthase II, 220 aa   ORF14: Coenzyme F390 synthetase, 484 aa
ORF28: O-methyltransferase, 350 aa   ORF13: O-acyltransferase, 378 aa

Fig. 3B

Fatty acid

⬇

ORF17: Epoxide hydrase
ORF20: Monooxygenase
ORF21: Iron-sulfur flavoprotein
ORF29: P-450 hydroxylase
ORF30: Oxidoreductase
ORF32: Oxidoreductase
ORF35: Proline oxidase
ORF38: P-450 hydroxylase

ORF13: O-acyltransferase, ORF19:Glycosyl transferase, ORF23: Type II NRPS condensation enzyme

Fig. 4

Fig. 5A
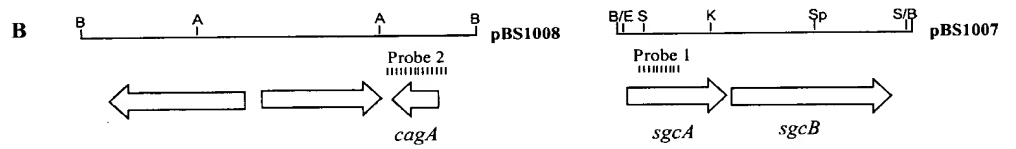
Fig. 5B

Fig. 5C

Fig. 6

```
        BamHI    EcoRI
1    GGATCCGGGAAGACCGGAATTCCGCCCCCAGCCCGGTCGAACTCGTATCGCTCCTGGTAGAACTGACGAAGCGTCATCGCCTGAC AAGGAGGCGGACCG   100
                                                                                            sgcA >

101  ATGAGGATGCTGGTGACGGGCGGAGCGGGTTTCATCGGCTGCGTCGTGCGGCGGCCACACTGCACGGCGAGCTGCCGGGTTCCGAGGACGCCCGGGTGA   200
     M  R  M  L  V  T  G  G  A  G  F  I  G  S  Q  F  V  R  A  T  L  H  G  E  L  P  G  S  E  D  A  R  V  T

201  CGGTCCTGGACAAGCTGACTTACTCCGGCAATCCGGCCAACCTCACCTCCGTCGCGGCCCATCCGCGGCGTACACCTTCGTCCAGGGCGACACCGTCGACCC   300
     V  L  D  K  L  T  Y  S  G  N  P  A  N  L  T  S  V  A  A  H  P  R  Y  T  F  V  Q  G  D  T  V  D  P
                                                       SacII

301  GCGCGGTCGTCGACGAGGTGGTCGCCGGCCACGACGTCATCGTCCACTTCGCGGCGGAGTCGCACGTGGACAGATCGATCGACACGGCCACCCGGTTCGTC   400
     R  V  V  V  D  E  V  V  A  G  H  D  V  I  V  H  F  A  A  E  S  H  V  D  R  S  I  D  T  A  T  R  F  V

401  ACGACCAACGTGCTCGGGAACGCTGCTGGAAGCGGCTCTCCGGAAGCGGGTTCGTCCACGTGTCGACGGACGAGGTCTACGGGTCGA   500
     T  T  N  V  L  G  T  Q  T  L  L  E  A  A  L  R  H  G  V  G  R  F  V  H  V  S  T  D  E  V  Y  G  S  I

501  TCGCCTCCGGCTCATGGACGGAGGACACCCCGCTCGCCCCCAACGTGCCCTACGGCGGTTCGGAAGGCGGTTCGACCGGCGCTGATGGCGCTGGCACCG   600
     A  S  G  S  W  T  E  D  T  P  L  A  P  N  V  P  Y  A  A  S  K  A  G  S  D  L  M  A  L  A  W  H  R

601  CACCCGGGGCCTGGACGTCGTCGTCACCCGTTGTACGGGCCGTACCAGTACCCGGAGAAGGTGATCCCGCTCTTCGTCACCAACATCCTC   700
     T  R  G  L  D  V  V  V  T  R  C  T  N  N  Y  G  P  Y  Q  Y  P  E  K  V  I  P  L  F  V  T  N  I  L

701  GACGGCTTGCGGGTGCCCCTGTACGGGGACGGCGCGCACCGGCGCGACTGGCTGCACGTGTCGGACCACTGCCGTGCTCACGGCGTGCGGCACCGACTGGTCCTG   800
     D  G  L  R  V  P  L  Y  G  D  G  A  H  R  R  D  W  L  H  V  S  D  H  C  R  A  I  Q  M  V  M  N  S  G

801  GCCGGGGGAGTCTACCACATCGGCGGCGGCCAGCGGCGACCGGCAGGGGGCCGACCGGGTTGGGCTACGAGCCCCTGGTCGCCTTC   900
     R  A  G  E  V  Y  H  I  G  G

901  CGTGGACCGGGTGGCCGACCGGCAGGGACGACCGGCGCGGGTAGACGGCGCCGCTACTCGCTGGACATCACGAAGATCCGTCAGGAGCTGGGCTACGAGCCCCTGGTCGCCTTC   1000
     V  D  R  V  A  D  R  Q  G  H  D  R  R  Y  S  L  D  I  T  K  I  R  Q  E  L  G  Y  E  P  L  V  A  F

1001 GAGGACGGCCTGGCCGCGACGGTGAAGTGGAAGTGGTACCACGAGAACCGTTCGTGGTGGCAGCCGCTCAAGGAGGCGGCGGGGCTGCTCGACGCCGTCGGCTGAC   1100
     E  D  G  L  A  A  T  V  K  W  K  W  Y  H  E  N  R  S  W  W  Q  P  L  K  E  A  A  G  L  L  D  A  V  G  *
                                       KpnI

1101 GGCAGCCACCGCTAGAGAACACCCCAG GAAAG GAGCCCACCTCCGTGACAGCAGTCAAGGAGCCCACGTCCCGCGCAGGACGCGGGAGTGGATCGCTCTCG   1200
     sgcB > M  T  A  V  K  E  P  T  S  R  A  G  R  R  E  W  I  A  L  V

1201 TCGTCCTCTCCTTGCCCACGATGCTGTTGATGCTGGACATCAACGTCCTCATGCTGGCCCTTGAGCCAGTTGAGCGAGCAGCACGCA   1300
     V  L  S  L  P  T  M  L  L  M  L  D  I  N  V  L  M  L  A  L  P  Q  L  S  E  D  L  G  A  S  S  T  Q

1301 ACAGCTGTGGATCACCGACATCTACGGATTCGCCATCGCCGGCTTCCTGGTGACCATGGGCACCCTCGGCGACCGGATCGGGCGCCGCCTCCTGCTC   1400
     Q  L  W  I  T  D  I  Y  G  F  A  I  A  G  F  L  V  T  M  G  T  L  G  D  R  I  G  R  R  L  L  L

1401 GGGGGCGGCGGCCGTGTTCGCCGTGTCCGTGGTCGCCGCCTTCTCCGACAGCGCCGCGATGCTCGTGTCAGCGCGTGCCGTCCTCGGCGTCGGCGGGG   1500
     G  G  A  A  V  F  A  V  V  S  V  V  A  A  F  S  D  S  A  A  M  L  V  V  S  R  A  V  L  G  V  A  G  A

1501 CCACGGTGATGCCCTCGACGCTCGCCATCAGCAACATGTTCGAGGACCCCAAGGAGCGCGGCACCGCCATCGCCATGTGGGCGAGCGCCATGATGGC   1600
     T  V  M  P  S  T  L  A  L  I  S  N  M  F  E  D  P  K  E  R  G  T  A  I  A  M  W  A  S  A  M  M  A

1601 CGGAGTCGCCCTCGGCCCGGCCGTCGGCGGGCTGGTGCTCGCCGCGTTCTGGTGGGGATCGGTGTTCCTCATCGCCGTTCCGGTGATGCTGCTGGTGGTG   1700
     G  V  A  L  G  P  A  V  G  G  L  V  L  A  A  F  W  W  G  S  V  F  L  I  A  V  P  V  M  L  L  V  V
```
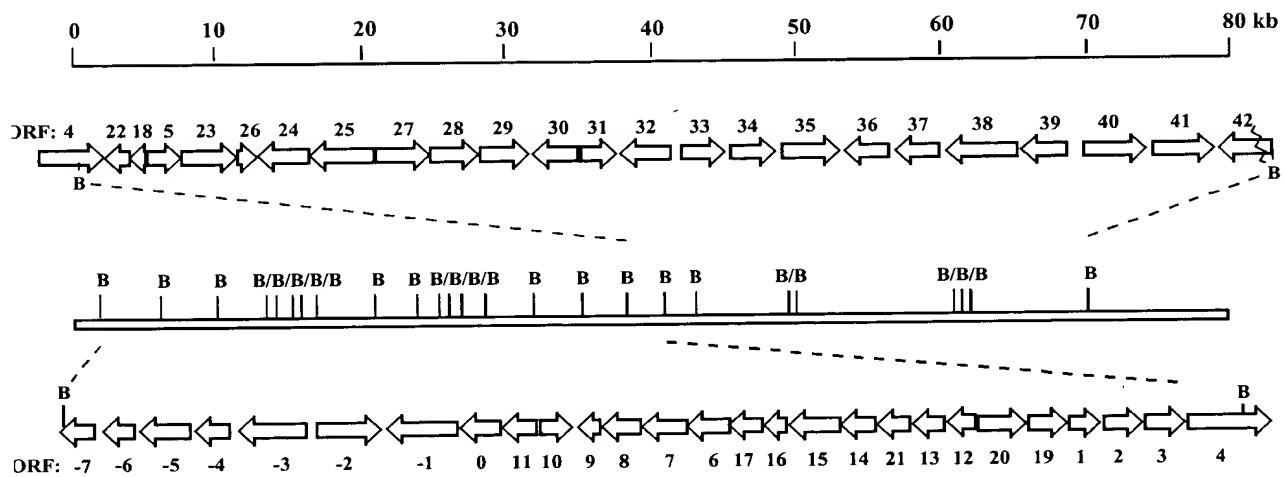
Fig. 6 cont'd.

```
1701  GTCACCGGCCCCGTGCTGCTCACCGAGTCCCGCGACCCCGACGCCGGACGGCCTGCTGAGCGCGGGGCTCTCCCTGCGCACCGTGCTGCCGGTGA  1800
       V  T  G  P  V  L  L  T  E  S  R  D  P  D  A  G  R  L  D  L  L  S  A  G  L  S  L  A  T  V  L  P  V  I

1801  TCTACGGACTGAAGGAGCTGGCCCGGACCGGGTGGGACCCGCTGGCGGCAGGTGCTGTCCTCGGCGTGATCTTCGGCGCGCTGTTCGTCCAGCGCCA  1900
       Y  G  L  K  E  L  A  R  T  G  W  D  P  L  A  A  G  A  V  V  L  G  V  I  F  G  A  L  F  V  Q  R  Q

1901  GCGGCGGTTGGCCGACCCCATGCTGGACCTCGGCCTCTTCGCCGACCGCACCCTGCGGGCGGGTCAGTGGTCATCGGC  2000
       R  R  L  A  D  P  M  L  D  L  G  L  F  A  D  R  T  L  R  A  G  L  T  V  S  L  V  N  A  V  I  M  G

2001  GGGACCGGACTGATGGTCGCCCTGTACCTCCAGACGATCGCCGGTCACTCCCCGTTGGCCGCGGCTGGCTGGCTGCTGATCCCGGCTGCATGCTCGTCG  2100
       G  T  G  L  M  V  A  L  Y  L  Q  T  I  A  G  H  S  P  L  A  A  G  L  W  L  L  I  P  A  C  M  L  V  V
                                                                                               SphI

2101  TGGGCGTACAGCTGTCGAACCTGCTGGCCCAGCGGATGCCCCCTTCCCGGGTGCTGCTGGGGGGACTGCTGATCGCCGCCGTCGGACAGCTCCTGATCAC  2200
       G  V  Q  L  S  N  L  L  A  Q  R  M  P  P  S  R  V  L  L  G  G  L  L  I  A  A  V  G  Q  L  L  I  T

2201  CCAGGTGGACACCGAGGAGACACCGCCCTCCATCGGCGCCCTGATCTACTTCGGCGCCGAGTTCGGACGGCCTGGGCATCGCGGGCCTGGGGAGTCATG  2300
       Q  V  D  T  E  D  T  A  L  L  I  A  A  T  T  L  I  Y  F  G  A  S  P  V  G  P  I  T  T  G  A  I  M

2301  GGAGCGCGCCGCCCCCGGAGAAGGCGGGGTCGGGGTGCCTCGTCGTGCCTCGTCGGCCGGCCGGCGACGCATCGCCGCCCTGCACACGGC  2400
       G  A  A  P  P  E  K  A  G  A  A  S  L  S  A  T  G  G  E  F  G  V  A  L  G  I  A  G  L  G  S  L  G

2401  GCACCGTCGTGTACAGCGCCGGGGTCGAGGTCGACGCGCCCTGGACTCCGGCGCGGCCGGCGCAGGAGAGCATCGCCGGGGCCCTGCACACGGCA  2500
       T  V  V  Y  S  A  G  V  E  V  P  D  A  A  G  P  A  D  A  D  A  A  Q  E  S  I  A  G  A  L  H  T  A

2501  CGGTCAGTGGCCACCGGGCCAGCGGGCAGCGGGACGCGGCCGCGGCCTTCACCAGCGGCGTGCAGTCGGTGTCGGTTC  2600
       G  Q  L  A  P  G  S  A  D  A  L  L  D  S  A  R  A  A  F  T  S  G  V  Q  S  V  A  A  V  C  A  V  F

2601  TCCCTGGCGCTCGCCGTCCTCATCGGCACCCGGCTGCGGGACATTTCGGCGATGGACCACGGGCACGGCGAGGAACGACGCTCAACCGG  2700
       S  L  A  L  A  V  L  I  G  T  R  L  R  D  I  S  A  M  D  H  G  H  G  E  E  P  A  E  N  D  A  Q  P  A

2701  CCACATGAGGCCGCACTTCCGGAGATGCAACGGCGTCCGAGTATGAGGATCAACGGCGTCCGTCGGAGTAGTGGAGTACT  2800
       T  *

2801  GGAACAGCACGGGCCGGAGACCATGCCCCGCGAGAACACTCACCTCCATGGCGGGCGACTACGGCGGCGGTGCCTCCTGCGCAAGGCCGACCTCCTGCGCAAGGCTTTCGCCCTT  2900
                                                                                                        SacII

2901  CTGGGCGGGAACGACTCCCCGAGAACATCACCTCCATGGCGGGCGACTACGGCGGCGGCGCGGGTGCCTCCTGCGCAAGGCCGACCTCCTGCGCAAGGCCGGGAAGCCGCG  3000
       BamHI

3001  TCTCCCCCTTACGGCACCTGGCCCTCGCTGGATCC  3035
```

Fig. 7

```
Gdh        1:~~~MRVLVTGGAGFIGSHYVRQLLGGAYPAFAGADVVVLDKLTYAGNEENLRPVADDPRF: 57
TylA2      1:~~~MRVLVTGGAGFIGSHFTGQLLTGAYPDLGATRTVVLDKLTYAGNPANLEHVAGHPDL: 57
SgcA       1:~~~MRMLVTGGAGFIGSQFVRATLHGELPGSEDARVTVLDKLTYSGNPANLTSVAAHPRY: 57
MtmE       1:MTTTSILVTGGAGFIGSHYVRTLLGPR..GVPDVTVTVLDKLTYAGTLTNLAEVSDSDRF: 58
consensus  1:   mrvLVTGGAGFIGShyvr lL g  pa     v VLDKLTYaGn  NL  Va   prf: 60


Gdh       58:RFVRGDICEWDVVSEVMREVDVVVHFAAETHVDRSILGASDFVVTNVVGTNTLLQGALAA:117
TylA2     58:EFVRGDIADHGWWRRLMEGVGLVVHFAAESHVDRSIESSEAFVRTNVEGTRVLLQAAVDA:117
SgcA      58:TFVQGDTVDPRVVDEVVAGHDVIVHFAAESHVDRSIDTATRFVTTNVLGTQTLLEAALRH:117
MtmE      59:RFVRGDICDAPLVDDLLAVHDQVVHFAAESHVDRSILGAADFVRTNVTGTQTLLDAALRQ:118
consensus 61:FVrGDi d  vv evm   dvvVHFAAEsHVDRSI  a  FV TNV GTntLL aAl  :120


Gdh      118:NVSKFVHVSTDEVYGTIEHGSWPEDHLLEPNSPYSAAKAGSDLIARAYHRTHGLPVCITR:177
TylA2    118:GVGRFVHISTDEVYGSIAEGSWPEDHPVAPNSPYAATKAASDLLALAYHRTYGLDVRVTR:177
SgcA     118:GVGRFVHVSTDEVYGSIASGSWTEDTPLAPNVPYAASKAGSDLMALAWHRTRGLDVVVTR:177
MtmE     119:GLETFVHISTDEVYGSIDAGSWPETAPVSPNSLYSAAKASSDLVALAYHRTHGLDVRVTR:178
consensus 121:gv kFVHvSTDEVYGsI  GSWpEd pl PNspY A KAgSDLiAlAyHRThGLdV vTR:180


Gdh      178:CSNNYGPYQFPEKVLPLFITNLMDGRRVPLYGDGLNVRDWLHVTDHCRGIQLVAESGRAG:237
TylA2    178:CSNNYGPRQYPEKAVPLFTTNLLDGLPVPLYGDGGNTREWLHVDDHCRGVALVGAGGRPG:237
SgcA     178:CTNNYGPYQYPEKVIPLFVTNILDGLRVPLYGDGAHRRDWLHVSDHCRAIQMVMNSGRAG:237
MtmE     179:CSNNYGSHQFPEKVIPLFVTSLLDGREVPLYGDGTNVRDWLHVDDHVRAIELVRTGGRAG:238
consensus 181:CsNNYGp QfPEKvlPLFiTnllDG  VPLYGDG n RdWLHV DHcRgi lV    GRaG:240


Gdh      238:EIYNIGGGTELTNKELTERVLELMGQDWSMVQPVTDRKGHDRRYSVDHTKISEELGYEPV:297
TylA2    238:VIYNIGGGTELTNAELTDRILELCGADRSALRRVADRPGHDRRYSVDTTKIREELGYAPR:297
SgcA     238:EVYHIGGGTELSNEELTGLLLTACGTDWSCVDRVADRQGHDRRYSLDITKIRQELGYEPL:297
MtmE     239:EVYNIGGGTELSNKELTQLLLDACGAGWDRVRYVTDRKGHDRRYSVDCTKIRRELGYRPA:298
consensus 241:eiYnIGGGTELtN ELT  vLe cG dws v  V DR GHDRRYSvD TKIr ELGY P :300


Gdh      298:VPFERGLAETIEWYRDNRAWWEPLKSAPDGGK~~~~:329
TylA2    298:TGITEGLAGTVAWYRDNRAWWEPLKRSPGGRELERA:333
SgcA     298:VAFEDGLAATVKWYHENRSWWQPLKEAAGLLDAVG~:332
MtmE     299:REFGDALAETVAWYRHHRAWWEPLTRAYGAVAA~~~:331
consensus 301:  f egLA Tv WYrdnRaWWePLk a gg        :336
```
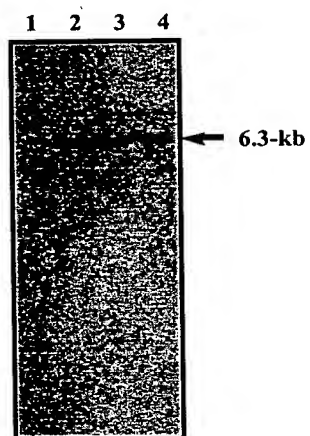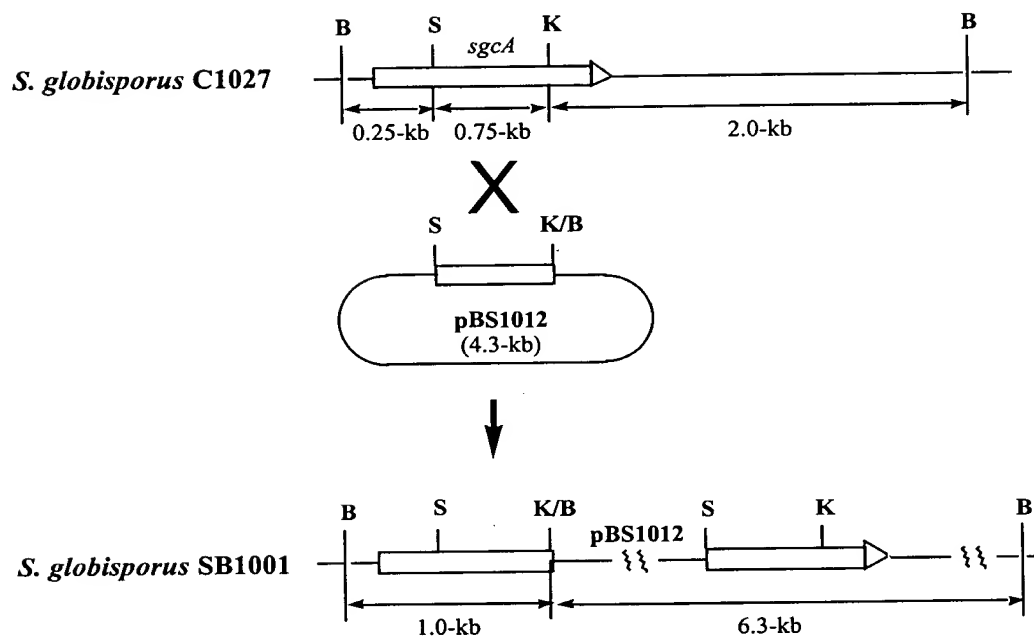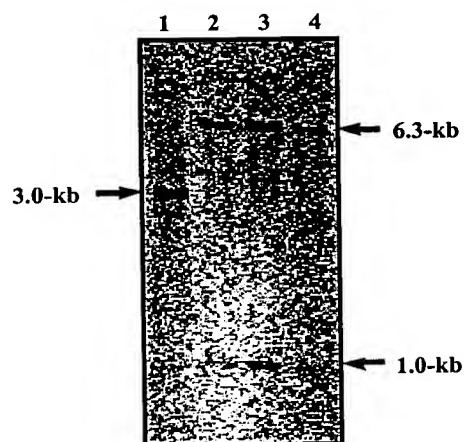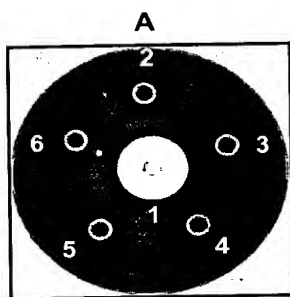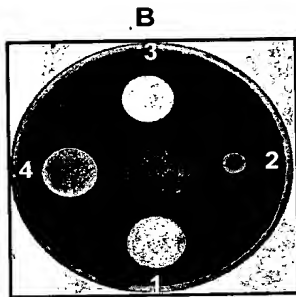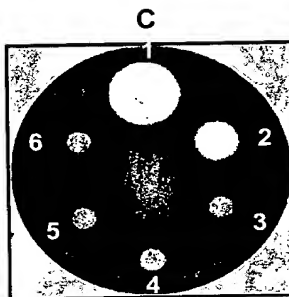
Fig. 8A

S. globisporus C1027

B   S   *sgcA*   K   B

0.25-kb   0.75-kb   2.0-kb

X

S   K/B

pBS1012
(4.3-kb)

S. globisporus SB1001

B   S   K/B   pBS1012   S   K   B

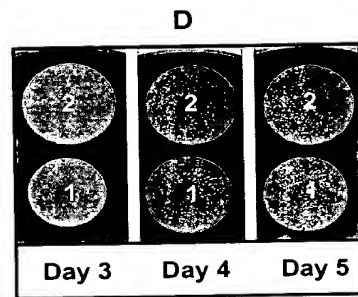1.0-kb   6.3-kb

6.3-kb

Fig. 8B

3.0-kb

6.3-kb

1.0-kb

Fig. 8C

-10-

Fig. 9A   **Fig. 9B**   **Fig. 9C**   **Fig. 9D**